MODULAR FORMS OF WEIGHT ONE: GALOIS REPRESENTATIONS AND DIMENSION.

NOTES DES EXPOSÉS 10, 12 ET 13 DU SÉMINAIRE FRANK THORNE.

DENIS TROTABAS

ABSTRACT. The present notes are the expanded and polished version of three lectures given in Stanford, concerning the analytic and arithmetic properties of weight one modular forms. The author tried to write them in a style accessible to non-analytically oriented number theoritists: in particular, some effort is made to be precise on statements involving uniformity in the parameters. On the other hand, another purpose was to provide an introduction, together with a set of references, consciously kept small, to the realm of Galois representations, for non-algebraists – like the author. The proofs are sketched, at best, but we tried to motivate the results, and to relate them to interesting conjectures. The author thanks warmly Daniel Bump, Brian Conrad, Kannan Soundararajan, Akshay Venkatesh and Justin Walker for helpful comments and inspiring conversations.

1. SOME NOTATIONS.

Let $\mathfrak{H} = \{z \in \mathbf{C}; \Im(z) > 0\}$ be the Poincaré upper-half plane.

Let k and q be two integers, and, as usual, $\Gamma_0(q)$ be the subgroup of $\mathrm{SL}(2, \mathbf{Z})$ of matrices whose lower left entries are divisible by q. It acts on $\mathfrak{H}$ by fractional linear transformations: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot z = \frac{az+b}{cz+d}$.

Let χ be a Dirichlet character modulo q: it defines a character on $\Gamma_0(q)$, by evaluating χ at the lower right entry.

If X is a finite set, $|X|$ denotes its cardinality; we reserve the letters p, ℓ for prime numbers, and n, m for integers; $\pi(x)$ is the number of primes less than x. Recall that $\pi(x) \underset{x \to \infty}{\sim} \frac{x}{\log(x)}$.

The letters K, E, k (resp. K_λ) denote fields (resp. the completion of K with respect to the valuation associated to λ), and $\mathcal{O}_K, \mathcal{O}_\lambda$ stand for the rings of integers of K, K_λ in the relevant situations.

For a complex number z, the notation $\mathrm{e}(z)$ stands for $\exp(2\pi i z)$.

The notation $f(x) \ll_A g(x)$ means that for any A, there exists a real number $C(A)$ such that for any x, $|f(x)| \leq C(A) \cdot |g(x)|$; if one adds "as $x \to \infty$", it means that the last inequality holds for $x \geq x(A)$ for some real number $x(A)$. In the same spirit, the notation $f(x) = o_{x \to x_0}(g(x))$

1

(resp. $f(x) = \mathcal{O}_{x \to x_0}(g(x))$) means that the quotient $f(x)/g(x)$ is defined in a (pointed) neighbourhood of x_0, and that $|f(x)/g(x)|$ tends to zero (resp. stays bounded) when x tends to x_0.

Sometimes, we used the notation $x \doteq y$ to mean $x = cy$ for some constant c, specifying the parameters on which c may depend; in these instances, the reader can check that the corresponding identities are not as trivial as this notation suggests.

2. Modular forms

2.1. For any holomorphic function f defined on $\mathfrak{H}$ and $\gamma \in \Gamma_0(q)$, we define:

$$f_{|\gamma}(z) = \chi(\gamma)^{-1}(cz + d)^{-k} f(\gamma(z))$$

Consider the following properties:
(M1): For any $\gamma \in \Gamma_0(q)$, $f_{|\gamma} = f$.
This implies, by Fourier analysis, that for any $\sigma \in \mathrm{SL}(2, \mathbf{Z})$ one has a decomposition:

$$f_{|\gamma}(z) = \sum_{n \in \mathbf{Z}} c_n(f, \sigma) \exp(2i\pi n z)$$

The holomorphy at $i\infty$ is then expressed by:
(M2): For any $\sigma \in \mathrm{SL}(2, \mathbf{Z})$, $c_n(f, \sigma) = 0$ for all negative n.
(M2'): For any $\sigma \in \mathrm{SL}(2, \mathbf{Z})$, $c_n(f, \sigma) = 0$ for all non-positive n.

2.2. The space of modular forms of weight k, level q and nebentypus χ is the set of holomorphic functions satisfying (M1) and (M2) above; the subspace of modular forms satisfying (M2') as well is called the space of cusp forms, noted $\mathcal{S}_k(q, \chi)$. The latter is finite dimensional (as is the first), and equipped with the Petersson inner product, invariant under the group action (it is a quotient of a Haar measure):

$$\langle f, g \rangle = \int_{\Gamma_0(q) \backslash \mathfrak{H}} f(x + iy) \overline{g(x + iy)} y^k \frac{dxdy}{y^2}$$

Note right now that by taking $\gamma = -\mathrm{I}$, (M1) gives $f(z) = (-1)^k \chi(-1) f(z)$, so if χ and k don't have the same parity, the space of modular forms is $\{0\}$; we shall exclude this case.

2.3. **Hecke operators.** On the space of modular forms of weight k and level q, one has the so-called Hecke operators, defined as follows for any prime p:

- If $(p, q) = 1$, $T_p(f)(z) = \sum_n c_{pn}(f)\mathrm{e}(nz) + \chi(p)p^{k-1}\sum_n c_n(f)\mathrm{e}(pnz)$ – p is called a good prime.
- If $p|q$, $T_p(f)(z) = \sum_n c_{pn}(f)\mathrm{e}(nz)$ – p is a bad prime.

The Hecke operators preserve the space of cusp forms; the Hecke operators at good primes all commute, and are normal with respect to the Petersson inner product. These important facts are explained in Miyake [M], as are the multiplicativity relations. In particular, if f is an eigenfunction for all the Hecke operators at good primes, with eigenvalues $\{a_p(f)\}$, one has $c_p(f) = c_f(1)a_p(f)$ at good p. To diagonalize further the Hecke operators, and get a good definition of L-series, it is necessary to introduce

2.4. **Newforms and oldforms.** Suppose χ defines a Dirichlet character modulo q', for $q'|q$. For any cusp form g in $\mathcal{S}_k(q', \chi)$, one checks easily that $z \mapsto g(dz)$ defines an element of $\mathcal{S}_k(q, \chi)$, for any $d|q/q'$. Let

$$\mathcal{S}_k^{old}(q, \chi) = \bigcup_{\substack{\chi \text{factors through } q'|q \\ d|q/q'}} \{z \mapsto g(dz) \; : \; g \in \mathcal{S}_k(q', \chi)\}$$

be the space of oldforms, and let

$$\mathcal{S}_k^{new}(q, \chi) = \mathcal{S}_k^{old}(q, \chi)^\perp$$

be the space of newforms (it may be zero!). Then it can be shown that the whole Hecke algebra (i.e. including bad primes) can be diagonalized on the space of newforms. The primitive Hecke eigenforms (those with $c_1(f) = 1$) have distinct eigenvalues outside a finite number of primes ("multiplicity one", well explained in the adelic setting by Casselman [C], cf. Gelbart [G] as well). Their L-series have an Euler product, absolutely convergent if $\Re(s) > 1 + k/2$:

$$L(s, f) := \sum_n \frac{a_n(f)}{n^s} = \prod_p L(s, f_p)$$

with

$$L(s, f_p) = \left(1 - a_p(f)p^{-s} + \chi(p)p^{k-1-2s}\right)^{-1}$$

$$= \left(1 - \alpha_1(p, f)p^{-s}\right)^{-1} \left(1 - \alpha_2(p, f)p^{-s}\right)^{-1}$$

at a good prime p, and

$$L(s, f_p) = \left(1 - a_p(f)p^{-s}\right)^{-1}$$

at a bad prime, along with an analytic continuation (easy to see with the Mellin transform), functional equation – cf. Bump [Bu], Miyake [M], Iwaniec [I], etc.

When one proves a theorem, one can often reduce it to the case of newforms, thanks to this decomposition; be aware it might be tacitly done in what follows...

3

2.5. **Ramanujan conjecture.** Let f be a primitive newform. The Ramanujan conjecture is the following inequality:

$$|a_p(f)| \leq 2p^{\frac{k-1}{2}}$$

which is equivalent to $|\alpha_i(p,f)| = p^{\frac{k-1}{2}}$. It has been a theorem for 35 years now, proven by Deligne for weight greater than two. We will prove below the case $k = 1$ (cf. corollary 5.1.1).

2.6. **Rationality properties.** Let $f \in \mathcal{S}_k(q,\chi)$ be a eigenform for all the Hecke operators at good primes, with Hecke eigenvalues $\{a_f(p)\}_{p \nmid q}$. Then:

$$\mathbf{Q}(f) := \mathbf{Q}(a_f(p), \chi(p) : p \nmid q)$$

is a finite extension of $\mathbf{Q}$, and all the Hecke eigenvalues are integers in this extension. If the nebentypus is trivial, then this extension in totally real. Serre explains this cohomologically in his Durham lectures.

2.7. An interesting problem is the evaluation of the dimension of the space of cusp forms, when one or more of the parameters (k,q) vary. For instance, using Eichler-Selberg trace formula one can prove that (see Knightly-Li [KL] theorem 29.5):

$$(1) \qquad \dim(\mathcal{S}_k(q,\chi)) = \frac{k-1}{12}\psi(q) + \mathcal{O}\left(q^{1/2}\tau(q)\right)$$

uniform in $k \geq 2$ and q, where $\psi(q) = q\prod_{p|q}(1+p^{-1})$ and $\tau(q)$ is the number of divisors of q.

Similarly, one can bound the dimension of the space of new-forms using the Petersson trace formula (Iwaniec-Luo-Sarnak [ILS]), and one has a uniform estimate for q squarefree, $k \geq 2$:

$$(2) \qquad \dim(\mathcal{S}_k^{new}(q,\chi)) = \frac{k-1}{12}\varphi(q) + \mathcal{O}\left((kq)^{2/3}\right)$$

with $\varphi(q) = q\prod_{p|q}(1-p^{-1})$ the Euler phi ([fi]) function (sometimes called [faj]).

2.8. We are interested in this note in the case $k = 1$, for which of course one cannot apply the above formulas. The first one "suggests" that the main term vanishes, and that we should expect an asymptotic of the size of $\sqrt{q}$. More precisely, the error term in (1) comes from a sum involving class numbers of quadratic fields: we will be surprised to see that this fact is preserved! One purpose of this note is to make these heuristics a bit more precise.

4

2.9. **Construction of dihedral weight one modular forms.** For simplicity, let's assume the level q is prime in this section only, and let $\chi_q = \left(\frac{\cdot}{q}\right)$ be the Legendre symbol (hence $q \equiv 3\,(4)$ is the parity condition stated above). One can construct cusp forms of arbitrary weight using induction of Hecke characters associated to quadratic fields; see Miyake [M] §4.8. Let $K_q = \mathbf{Q}(\sqrt{-q})$, and let ψ be a character of the class group of K. Let:

$$\text{(3)} \qquad \theta_\psi(z) := \sum_{\mathfrak{n} \subset \mathcal{O}_K} \psi(\mathfrak{n}) \mathrm{e}(N(\mathfrak{n})z)$$

Then θ_ψ is actually a weight one primitive cusp form with nebentypus χ_q, unless ψ is real (in which case, the theta series is a weight one Eisenstein series). They are paired two by two, and one gets $\frac{h(K)-1}{2}$ independent such forms. The construction can be carried out for ray class groups, and for real quadratic fields as well (but in the last case, one gets a non-squarefree level).

Let's call $\mathcal{S}_1^{Dih}(q,\chi)$ the span of these theta series; by Siegel's theorem (for the lower bound):

$$q^{1/2-\varepsilon} \ll_\varepsilon \dim\left(\mathcal{S}_1^{Dih}(q,\chi)\right) = \frac{h(K)-1}{2} \ll_\varepsilon q^{1/2+\varepsilon}$$

which gives the $\sqrt{q}$ term promised earlier. The big task is to understand what are the forms we missed in this construction, and why they should be rare (though it is still conjectured). For the moment, let me state the first form of:

CONJECTURE 2.1. *One has the estimate, for q varying among the squarefree integers:*

$$\dim\left(\mathcal{S}_1(q,\chi)\right) = \frac{h(K_q)}{2} + \mathcal{O}_\varepsilon(q^\varepsilon)$$

with an $\mathcal{O}$-constant independent of q,χ.

2.10. **Remark.** From the point of view of representation theory, the theta series is really "induction" in the usual sense. Indeed, ψ defines a character of the Galois group $G_K = Gal(\overline{K}/K)$ by class field theory, hence a one-dimensional representation. So its induction $\mathrm{Ind}_{G_K}^{G_\mathbf{Q}}\psi$ defines a two-dimensional complex Galois representation. It is conjectured by Langlands that such operations on the Galois side still correspond to automorphic forms: in this case, the theta series *is* the corresponding form. For more on this issue, the paper of Rogawski [Rog] is excellent; Gelbart [G] sketches the proof in general for quadratic induction; Bump [Bu] has a very clear introduction to Langlands conjectures; Miyake [M] gives a proof in the classical setting; etc.

2.11. **How to tackle the conjecture?** By using the trace formula, the best estimate of the dimension of weight one modular forms is about q,

perhaps with a power saving in $\log(q)$. Duke achieved a power saving in q, by using two conflicting properties:

- "Smoothness" via harmonic analysis: a general feature of automorphic forms, a bit degenerated in weight one.
- "Rigidity" via Galois representations, imposing strong limitations.

The bound Duke gets is roughly of size $q^{11/12}$, using the full strength of Serre's analysis of the possible lifts from PGL(2) to GL(2); if one is happy with a tiny power saving, one can soften the input a lot.

3. Harmonic analysis

I will be brief here: it is difficult to give a better survey than Michel's Park City lecture notes [Mi]. Plus, I need only the large sieve inequality – which is the small residue of harmonic analysis left in weight one. The book of Iwaniec and Kowalski [IK] contains many examples of such inequalities as well, and explains how one can guess the "best" bound.

3.1. Let $\mathcal{B}_1(q,\chi)$ be an *orthonormal* basis of $\mathcal{S}_1(q,\chi)$. Recall that $c_n(f)$ denotes the n-th Fourier coefficient of such a form.

PROPOSITION 3.1. *Let $\{x_n\}_{n\in\mathbf{N}}$ be an arbitrary sequence of complex numbers. The following estimate is uniform in N, q, χ:*

$$(4) \qquad \sum_{f\in\mathcal{B}_1(q,\chi)} \left| \sum_{n=1}^{N} x_n c_n(f) \right|^2 \ll \left(1 + \frac{N}{q}\right) \sum_{n=1}^{N} |x_n|^2$$

The estimate (4) is called *large sieve inequality*: it is a powerful substitute for Cauchy-Schwarz inequality. Three proofs are given in Michel's lectures: using Kuznetsov trace formula, Rankin-Selberg L-functions, and a variant of Iwaniec (explained in Duke's paper as well). The first one usually gives the best bounds, but requires that the family one averages over be spectrally complete. Instead of reproducing these proofs here, let's explain this spectral aspect.

3.2. The proof of (4) should start as follows: one expands the product on the left hand side; one needs then to estimate the average of the product of two Fourier coefficients over an orthonormal basis. Such a formula is a "Petersson-Kuznetsov"-type trace formula. For example, for weights greater than 2, one has (the implied constant depends on k and can be made explicit) the Petersson formula:

$$\sum_{f\in\mathcal{B}_k(q,\chi)} c_n(f)\overline{c_m(f)} \doteq \delta_{n,m} + \sum_{c\equiv 0\,(q)} \frac{K\ell_\chi(n,m,q)}{|c|} J_{k-1}\left(\frac{4\pi\sqrt{nm}}{|c|}\right)$$

with a Bessel function J_{k-1} and a twisted Kloosterman sum $K\ell_\chi$, which is an explicit exponential sum, and can be evaluated for example using Weil's bound.

Now let $\mathcal{M}_k(q,\chi)$ be an orthonormal basis of the space of cuspidal Maass cusp forms of weight k, nebentypus χ, and Φ a "nice" holomorphic, rapidly decreasing test function, one has the Kuznetsov formula:

$$\sum_{f \in \mathcal{M}_k(q,\chi)} \Phi(t_f) c_n(f) \overline{c_m(f)} + (\text{Eis.}) \doteq \delta_{n,m} + \sum_{c \equiv 0 \, (q)} \frac{K\ell_\chi(n,m,q)}{|c|} \widehat{\Phi}\left(\frac{4\pi\sqrt{nm}}{|c|}\right)$$

where: $1/4 + t_f^2$ is the Laplace eigenvalue of f, $\widehat{\Phi}$ is a Bessel-transform of Φ, (Eis.) denotes the contribution of the continuous spectrum parametrized by the Eisenstein series (for the definition fo Maass forms, their multipliers, etc., see Bump [Bu] or Michel [Mi]). To ensure the space of weight k Maass forms is not trivial, there is a parity-type condition on k, χ. Interestingly, holomorphic forms can be embedded into the space of Maass form, using:

$$f \in \mathcal{S}_k(q,\chi) \longmapsto y^{k/2} f(z).$$

It turns out that for $k \geq 2$, the image of this map is precisely the space of forms with Laplace eigenvalue $\frac{k}{2}\left(1 - \frac{k}{2}\right)$, which is an isolated point of the spectrum of the Laplacian. By choosing a relevant test function (i.e. taking Φ with support outside the vertical line parametrized by the continuous spectrum), one can recover Petersson formula.

For $k = 1$, $1/4$ is no longer away from the continuous spectrum, but one can still prove a "large sieve inequality" for weight one holomorphic forms using such an embedding, and a positivity argument (a sum of positive numbers on $\mathcal{B}_1(q,\chi)$ can certainly be bounded by such a sum on $\mathcal{M}_1(q,\chi)$!). The first impression is that such a naive argument loses a lot of information. However, by re-establishing good harmonic behaviour, one can actually improve Duke's bound, without changing the input: I will come back to this later.

3.3. **Remark.** If χ is an even $(\chi(-1) = 1)$ character, the non-holomorphic Maass cusp forms with eigenvalue $1/4$ are conjectured to correspond to Galois representations, as do their "odd" cousins, the holomorphic weight one modular forms.

3.4. **Rankin-Selberg convolution.** Let f, g be two modular forms with the same weight (the construction works for any weight, but requires a slightly different normalization), possibly with distinct levels and nebentypus. We saw how to define a reasonable theory of L-functions, so let's suppose that f, g are both primitive newforms (or at least Hecke eigenforms at the good primes). One then defines (absolute convergence for $\Re(s) \gg 0$):

$$L(s, f \times \bar{g}) = \zeta(\chi_f \overline{\chi_g}, 2s) \sum_{n \geq 1} \frac{a_n(f) \overline{a_n(g)}}{n^s}$$

7

This L-function satisfies a functional equation, has a simple pole at $s = k$ if and only if $f = g$ (up to a scalar multiple), and has an Euler product, whose local factors at good primes are given by:

$$L(s, f_p \times \overline{g_p}) = \prod_{(i,j) \in \{1,-1\}} (1 - \alpha_i(p, f)\overline{\alpha_j(p, g)}p^{-s})^{-1}.$$

As $L(s, f \times \bar{f})$ has non-negative coefficients, Landau's lemma implies the convergence of the series in the region $\{\Re(s) > k\}$, from which one can infer that $|\alpha_i(p, f)| < p^{k/2}$ (by the absence of poles in this region applied to the factor at p) – and many more important results, like $\sum_{n=1}^{N} |c_n(f)|^2 \ll_{\varepsilon, f} N^{k+\varepsilon}$ for any $\varepsilon > 0$, which is Ramanujan on average, cf. Michel on this issue.

In the region of absolute convergence (and ignoring the bad primes, finite in number)

$$\log\left(L(s, f \times \bar{f})\right) = \sum_{m=1}^{\infty} g_m(s)$$

where

$$g_m(s) = - \sum_{(p,q)=1} \sum_{(i,j) \in \{1,-1\}} \frac{\alpha_i(p, f)^m \overline{\alpha_j(p, f)}^m p^{-ms}}{m}$$

one has

$$\sum_{m=1}^{\infty} g_m(s) = -\log(s - k) + \mathcal{O}_{s \to k}(1)$$

and so

$$g_1(s) = \sum_{(p,q)=1} |a_p(f)|^2 p^{-s} < -\log(s - k) \text{ for } s \to k$$

3.5. **Remark.** Once the Ramanujan conjecture is proven, it is easy to see that $\sum_{n \geq 2} g_m(s)$ stays bounded at 1, so the above inequality is actually an asymptotic.

3.6. **Consequence.** Suppose f is a weight one modular form and an eigenvector of the good Hecke operators. Then, the Hecke eigenvalues are integers of a finite extension $\mathbf{Q}(f)$ of $\mathbf{Q}$, so for any real number M the set:

$$S(M) = \{\alpha \in \mathcal{O}_{\mathbf{Q}(f)} : |\sigma(\alpha)| \leq M \, \forall \sigma : \mathbf{Q}(f) \hookrightarrow \mathbf{C}\}$$

is finite. So $\{a_f(p) : p \nmid q \text{ and } a_f(p) \in S(M)\}$ is finite. Also one has for $\Re(s) > 1$:

$$\sum_{p \nmid q} \frac{|a_f(p)|^2}{p^s} \geq \sum_{\substack{p \nmid q \\ a_f(p) \notin S(M)}} \frac{M^2}{p^s}$$

8

and so:

$$\sum_{\substack{p \nmid q \\ a_f(p) \notin S(M)}} p^{-s} < -M^{-2} \log(s-1)$$

which, by the definition of the (upper) Dirichlet density, means:

$$\mathrm{Dens}(p \nmid q : a_f(p) \notin S(M)) \leq M^{-2}$$

and we can summarize:

LEMMA 3.1. *For any positive η, there exists a finite subset Y_η of $\mathbf{C}$, and a set of primes $\mathscr{P}_\eta$ of density less than η such that:*

$$\forall p \notin \mathscr{P}_\eta \,, \, a_f(p) \in Y_\eta.$$

This means that, possibly outside a small set of primes, the Hecke eigenvalues of weight one modular forms (with fixed weight, nebentypus) are finite in number. Deligne-Serre rules out this possible tiny set.

3.7. **Densities.** A subset $\mathscr{P}$ of prime numbers has upper Dirichlet density α if:

$$\limsup_{s \to 1} \frac{\sum_{p \in \mathscr{P}} p^{-s}}{\sum_{p \text{ prime}} p^{-s}} = \alpha.$$

If the above quotient actually has a limit, one talks about "Dirichlet density". The (upper) natural density would be defined by:

$$\limsup_{x \to \infty} \frac{1}{\pi(x)} \{p \in \mathscr{P}, p \leq x\}$$

again with the symbol "lim" when it makes sense. Its existence implies the existence of Dirichlet density. For integers, one has of course the two corresponding definitions, again with the same comparisons. We will always write $\mathrm{Dens}(\mathscr{P})$ to denote the (upper) Dirichlet or natural density; as soon as the Tchebotarev density theorem is involved, natural densities actually exist.

4. GALOIS REPRESENTATIONS.

Conseils de lecture: I already mentionned Rogawski's review [Rog] on Artin conjecture , Langlands-Tunnell theorem. I should add Taylor's article [T], very complete but more difficult. Bump [Bu] has in his first chapter an excellent introduction to the topic, and it is not possible to me to forget Rohrlich's paper [Roh] on the Weil-Deligne group and elliptic curves. Many other things are to be found in Cornell-Silverman-Stevens [CSS]. Bushnell and Henniart's book [BH] on the local Langlands conjectures is also a good companion.

Let k be a complete field (topologized with an absolute value, allowed to be trivial only if k is a finite field), V a finite dimensional topological vector space on k (so topologically $V \cong k^n$ for some n). Let K be a field.

A representation of the absolute Galois group $Gal(\overline{K}/K) = G_K$ in V is a pair (ρ, V), often denoted ρ, where V is as above, and ρ is a continuous homomorphism:

$$\rho : G_K \longrightarrow \mathrm{GL}(V).$$

• ρ is *irreducible* if there are no closed subspace of V invariant under ρ.
• ρ is *semi-simple* if V is a direct sum of irreducible closed subspaces.
We will be interested in the cases where k is $\mathbf{C}$, a $\mathfrak{p}$-adic field or a finite field, with its natural topology.

4.1. **Remark.** Choose an algebraic closure $\bar{k}$ of k, and suppose that there are countably many finite extensions of k (inside this choice of $\bar{k}$). Then I claim that there exists a finite extension E of k such that $\rho(G_K) \subset \mathrm{GL}_n(E)$ – i.e. ρ is rational. Indeed, $\rho(G_K)$ is compact, and one can write

$$\rho(G_K) = \bigcup_{[E:k]<\infty} \left(\rho(G_K) \cap \mathrm{GL}_n(E) \right).$$

By the Baire category theorem (E, as a finite dimensional k-vector space, is endowed with the product topology hence is complete), one of the $\rho(G_K) \cap \mathrm{GL}_n(E)$ has non-empty interior in $\rho(G_K)$, so is cofinite inside it; by adding the remaining matrices, one gets the claim.

4.2. **The semi-simplicity issue.** G_K is always compact. So if $k = \mathbf{C}$, any representation is semi-simple (one can define the topology of V by a G_K-invariant product, and can thus take orthogonal complements; this fact extends of course in the case of Hilbert spaces).
In the other cases, it is a bit more complicated. But if G is a finite group, then any representation of G is semisimple, as long as $\mathrm{char}(k)$ is prime to $|G|$; in particular it is always true for $\mathfrak{p}$-adic fields.
The importance of semi-simple representations is that they are determined (up to isomorphism) by their character, or characteristic polynomial:

PROPOSITION 4.1. *Let G be a group, ρ, ρ' two k-semi-simple representations. If for any $g \in G$, $\mathrm{Tr}(\rho)(g) = \mathrm{Tr}(\rho')(g)$ for $\mathrm{char}(k) = 0$ (resp. $\det(Id - X\rho(g)) = \det(Id - X\rho'(g))$ if k is finite), then ρ, ρ' are isomorphic.*

In the finite case, the proof can be found in Curtis-Reiner [CR] (30.16) e.g. Look in Husemöller's elliptic curves [H] §15.2 for some comments in the ℓ-adic case.

If ρ is not semi-simple, then one can define from ρ a semi-simple representation (unique up to isomorphism). Indeed, let $\{V_i\}$ be a finite Jordan-Hölder composition series of V: $\{0\} = V_0 \subset V_1 \subset \cdots \subset V_N = V$ where all inclusions are strict, and V_i is a maximal G-submodule of V_{i+1}. Consider $\mathrm{Gr}(V) = \bigoplus_{i \geq 0} V_{i+1}/V_i$, which carries a G-action, induced from the one on G: this representation is semi-simple by construction, and has the same character, characteristic polynomials, as ρ. It is called *the* semi-simplification of ρ (by the "uniqueness" implied by the last proposition).

10

4.3. **Complex case.** $G_{\mathbf{Q}}$ has a basis of neighbourhoods of the identity consisting of subgroups, namely $\{Gal(\overline{\mathbf{Q}}/E)\}_{[E:\mathbf{Q}]<\infty}$. But $\mathrm{GL}_n(\mathbf{C})$, as any Lie group, has a neighbourhood of the identity which doesn't contain any subgroup: this is easily seen by looking at det : $\mathrm{GL}_n(\mathbf{C}) \to \mathbf{C}^{\times}$, as if $U \subset \mathrm{GL}_n(\mathbf{C})$ is small enough, $\det(U)$ is contained in a small disc centered at 1, which cannot contain any subgroup of $\mathbf{C}^{\times}$. If U is chosen this way, then for $E/\mathbf{Q}$ large enough, $\rho(Gal(\overline{\mathbf{Q}}/E)) \subset U$, so is trivial! As a consequence, any complex Galois representation factors through some finite Galois extension E of $\mathbf{Q}$, which can of course be chosen so that $\rho : Gal(E/\mathbf{Q}) \to \mathrm{GL}_n(\mathbf{C})$ is injective.

4.4. **If $K = \mathbf{Q}$.** Let (ρ, V) be a complex or $\mathfrak{p}$-adic Galois representation. Let us fix once and for all an algebraic closure of $\mathbf{Q}, \mathbf{Q}_p$ for all the primes p, and an embedding $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}_p}$ for each p. From these embeddings, one deduces a system of local representations:

$$\rho_p : G_{\mathbf{Q}_p} \longrightarrow \mathrm{GL}(V).$$

Recall the short exact sequence

$$\{1\} \longrightarrow I_p \longrightarrow G_{\mathbf{Q}_p} \longrightarrow G_{\mathbf{F}_p} = \widehat{\mathbf{Z}} \longrightarrow \{0\}$$

where $\widehat{\mathbf{Z}}$ is topologically generated by the Frobenius $x \mapsto x^p$. One can lift this Frobenius to an element of $G_{\mathbf{Q}_p}$, defined up to the inertia subgroup I_p. Let's fix a choice for each p of such a Frobenius element noted F_p. It defines an operator on inertia-invariant vectors V^{I_p}, and so one can define, if $k = \mathbf{C}$

$$L(s, \rho_p) := \det\left(Id - X\rho(F_p)\big|V^{I_p} \right)^{-1}\Big|_{X=p^{-s}}$$

It is important to remark that this L-factor does not depend on the choice of $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}_p}$.

If k is $\mathfrak{p}$-adic, one has to choose first an algebraic embedding $\imath : k \hookrightarrow \mathbf{C}$, and then take

$$L(s, \rho_p) := \imath\left(\det\left(Id - X\rho(F_p)\big|V^{I_p} \right)^{-1} \right)\Big|_{X=p^{-s}}$$

Note that the last L-factor depends on the algebraic embedding $\imath$, but fortunately Deligne tells us that in geometric situations, the relevant determinant lies in a number field.

The global L-function, called the Artin L-function, is defined for $\Re(s) \gg 0$ by

$$L(s, \rho) = \prod_p L(s, \rho_p)$$

Actually, if $k = \mathbf{C}$, the eigenvalues of the Frobenius elements have modulus one, so it is easy to check the covergence for $\Re(s) > 1$; in the $\mathfrak{p}$-adic case, I don't know in general: if ρ arises in the étale cohomology of an algebraic

11

variety, then Deligne's theorem (Weil I or II) gives a bound on the eigenvalues of the geometric Frobenius elements, from which one can deduce the convergence.

4.5. **Remark.** In the complex case, ρ factors through a finite quotient, which already kills almost all inertia subgroups; one deduces that almost all the local L-factors have degree n in p^{-s} – one says that ρ is *unramified* at p. In the $\mathfrak{p}$-adic case, it can be false in general; but if ρ is geometric, then Grothendieck *et al.* showed that ρ actually factors through a quotient $Gal(E/\mathbf{Q})$, usually infinite, such that $E/\mathbf{Q}$ is unramified outside a finite number of primes. This is nice, because one can use a powerful tool: the Tchebotarev theorem.

4.6. **Tchebotarev density theorem.** It makes possible a refinement of proposition 4.1. Let's state it in the most general form:

THEOREM 4.1. *Let $E/\mathbf{Q}$ be an algebraic Galois extension, unramified outside a finite set of primes. Let C be a measurable subset of $Gal(E/\mathbf{Q})$ stable under conjugation, with* $\mathrm{vol}(\partial C) = 0$. *Then:*

$$\mathrm{Dens}\left(p \ unramified \ prime \ : \ F_p \subset C \right) = \frac{\mathrm{vol}(C)}{\mathrm{vol}(Gal(E/\mathbf{Q}))}$$

Here "density" means natural or Dirichlet density; note that in this statement, one considers the Frobenius elements as conjugacy classes, as one did not fix a priori any embedding $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}_p}$: as a consequence, one has to take into account the action of $Gal(E/\mathbf{Q})$ on the decomposition group (by conjugation; a transitive action), and the Frobenius only defines a conjugacy class. We'll talk about "finite" (resp. "infinite") Tchebotarev if $E/\mathbf{Q}$ is a finite (resp. infinite) extension.

Let $g \in Gal(E/\mathbf{Q})$. For any finite Galois subextension $K/\mathbf{Q}$, one can apply Tchebotarev to the conjugacy class of the reduction of g modulo $Gal(K/\mathbf{Q})$: it tells that this reduction g_K belongs to the conjugacy class F_p for infinitely many primes p. But one the other hand, $g_K \to g^{\dagger}$, so *the Frobenius classes are dense in $Gal(E/\mathbf{Q})$.* Another way to deduce this from the above version of Tchebotarev, is to take $C = \{hUh^{-1} \ : \ h \in G_{\mathbf{Q}}\}$ for an open neighbourhood of g of the type $U = g \cdot Gal(E/K)$. In any case, one deduces the sought amelioration of prop. 4.1:

PROPOSITION 4.2. *Suppose k is $\mathbf{C}$, a finite or a $\mathfrak{p}$-adic field. Let ρ, ρ' be two k-semi-simple representations of $G_{\mathbf{Q}}$. If for any unramified p, oustide a zero-density set of primes,* $\mathrm{Tr}\big(\rho(F_p)\big) = \mathrm{Tr}\big(\rho'(F_p)\big)$ *for* $\mathrm{char}(k) = 0$ *(resp.* $\det\left(Id - X\rho(F_p) \right) = \det\left(Id - X\rho'(F_p) \right)$ *if k is finite), then ρ, ρ' are isomorphic.*

---

†to make this convergence meaningful, consider the filter of the sections for the order defined by the inclusion of subextensions.

The case where k is a finite field case is not completely justified, but here again the Galois representation factors through a finite quotient, and one needs only the "finite" Tchebotarev theorem.

4.7. **Remark.** It is not easy to find a reference with a proof of the above version, which I stated for purely aesthetic reasons, as the condition on the boundary of C is quite restrictive in practice; see for example Serre [S2]. One can actually deduce it from the "finite" version, amply covered in the literature. Indeed, the "finite" version implies that for any subset C stable by conjugation,

$$\mathrm{Dens}\Big(p \text{ unramified prime } : F_p \subset CGal(E/K)\Big) = \frac{|\pi_K(C)|}{|Gal(K/\mathbf{Q})|}$$

for any finite (Galois) subextension $K/\mathbf{Q}$, where we noted $\pi_K : Gal(E/\mathbf{Q}) \twoheadrightarrow Gal(K/\mathbf{Q})$ the projection. Then, as the Haar measure of the (infinite) Galois group is the projective limit of the measures of its finite quotients, one sees easily that

$$\mathrm{vol}(\overline{C}) = \lim_K \frac{|\pi_K(C)|}{|Gal(K/\mathbf{Q})|}$$

hence the condition on the boundary of C. But to actually justify that the primes in the theorem have a density, it is necessary to argue as follows. Let $Gal(E/\mathbf{Q})^{\#}$ be the set of conjugacy classes of $Gal(E/\mathbf{Q})$. Let $\pi : Gal(E/\mathbf{Q}) \twoheadrightarrow Gal(E/\mathbf{Q})^{\#}$ the (open) projection. The Haar measure defines a measure on the conjugacy classes (by pushforward); the "finite" Tchebotarev, together with the above fact on projective limits of measures, implies that for any $K/\mathbf{Q}$ finite, and any $c \in Gal(E/\mathbf{Q})^{\#}$

$$\frac{\log(X)}{X} \sum_{p \leq X} \delta_{F_p}(\mathrm{Char}_{cGal(E/K)}) \underset{x \to \infty}{\longrightarrow} \int_{Gal(E/\mathbf{Q})^{\#}} \mathrm{Char}_{cGal(E/K)}(g) d^{\#}g$$

But as the characteristic functions $\mathrm{Char}_{cGal(E/K)}$ are dense (in the space of continuous functions on $Gal(E/\mathbf{Q})^{\#}$), the above convergence is true for any continuous map (uniform boundedness theorem), hence the sequence of the Frobenius elements $\{F_p\}_p$ is equidistributed in $Gal(E/K)^{\#}$. From this, as one can approximate the characterisitic function of any measurable set C in $Gal(E/\mathbf{Q})^{\#}$ – with negligible boundary – by continuous functions (here the approximation is in the sense of the order on the space of real functions, not in norm, cf. Billingsley [B] theorem 2.1, or Serre [S2] appendix of chapter I), one gets the result.
The point is that the densities (natural or Dirichlet) on prime numbers are not a measure, and the naive argument ("take the intersection of the sets of primes") doesn't work; but one gets the result by equidistribution, in two steps: from special characteristic functions to continuous functions (by approximation in norm), then from continuous functions to more general characteristic functions (thanks to the order on real numbers).

4.8.   If $\rho : G_{\mathbf{Q}} \longrightarrow GL(V)$ is a complex or $\mathfrak{p}$-adic semi-simple Galois representation, we defined above its Dirichlet L-series by a convergent (in some half plane) Euler product. It turns out that this L-function satisfies a functional equation, of the type:

$$\Lambda(s,\rho) = q^{\frac{1-2s}{2}} \omega_\rho \Lambda(1-s,\rho^\vee)$$

where Λ is the product of L with some Γ-factors, describing the behaviour of ρ at infinity (cf. the Durham volume, Martinet (same volume as [S1]), and Tate in the Corvallis volume [Ta]); $\rho^\vee$ is the contragredient of ρ, ω_ρ is a complex number of modulus one. Brauer showed that this L-function extends to a meromorphic function on $\mathbf{C}$, and Artin conjectured that Λ is holomorphic unless ρ contains the trivial representation: cf. Heilbronn in Cassels-Fröhlich [CF] for the meromorphic continuation.

4.9. **Remark.** There is at least one other type of geometric L-function one should mention: the L-function obtained from algebraic varieties by "counting points" modulo a prime. For example, for abelian varieties, it turns out that this zeta function agrees with the L-function of the Galois action on the (dual of the) Tate module: cf. Rohrlich [Roh] for a proof of this in the case of elliptic curves, and much more.

4.10. **Remark.** Unlike complex Galois representations, usually $\mathfrak{p}$-adic Galois representations don't have finite image.

5. The Deligne-Serre theorem

5.1. **Statement of the theorem.**

THEOREM 5.1. *Let f be a weight one cuspidal modular form, with nebentypus χ, and level q, a Hecke eigenvector at the good primes. Then there exists an irreducible complex Galois representation, unique up to isomorphism, $\rho_f : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}(2, \mathbf{C})$ such that at any good prime p:*

$$L(s, \rho_p) = L(s, f_p).$$

COROLLARY 5.1.1 (The Ramanujan Conjecture). *Under the same hypotheses, at good primes one has:*

$$|a_p(f)| \leq 2$$

If furthermore f is a newform, one has:

$$|a_n(f)| \leq \tau(n) \ll_\varepsilon n^\varepsilon.$$

If f is any weight one modular form, then:

$$|c_n(f)| \ll_{\varepsilon,f} n^\varepsilon.$$

COROLLARY 5.1.2. *Under the same hypotheses, if f is a primitive newform, the local L factors actually agree on the ramified primes as well, the conductors of the Artin L-function and of the newform f agree, as do the ε factors.*

14

PROOF of Corollary 5.1.1: Firstly $a_p(f) = \text{Tr}(\rho_f(F_p))$; as $\rho_f(G_\mathbf{Q})$ is finite, the eigenvalues of $\rho_f(F_p))$ are roots of unity, hence the first point; for the second, note that by multiplicativity $a_n(f) = \prod_p a_{p^{v_p(n)}}(f)$, and finally $a_{p^k}(f) = \prod_{i+j=k} \alpha_1(p,f)^i \alpha_2(p,f)^j$ whence the second bound. The third is trivial from the decomposition fo the total space of modular forms into newforms, and oldforms. Q.E.D.

See Deligne-Serre for the proof of the second claim: by the theorem, the two L-functions agree except possibly on finitely many ramified primes; so the quotient of the L-functions is a finite product. Using the functional equation, one proves it easily by reductio ad absurdum.

The proof of the Ramanujan conjecture is much more difficult in weight $k \geq 2$, as the geometric nature of the Galois representation attached to f is needed to prove it, via the study of the weights in ℓ-adic cohomology (Deligne, Weil I).

5.2. **Langlands' conjectures.** These predict that Galois representations should "correspond" to certain adelic automorphic forms. The Deligne-Serre theorem belongs to this program. Recently, Khare *et al.* proved that reciprocally, any odd Galois representation is associated to a weight one modular form (odd means that the Dirichlet character χ, corresponding to $\det(\rho(.))$ by class field theory, satisfies $\chi(-1) = -1$), completing the Langlands-Tunnell theorem (cf. Rogawski [Rog]). With a little more work, one could strengthen the theorem with Langlands' philosophy in mind. If f is a newform, one can lift f to a square-integrable (modulo the center) function on $\text{GL}(2,\mathbf{Q})\backslash\text{GL}(2,\mathbb{A}_\mathbf{Q})$; let $V(f) = \overline{\text{span}\{f(\cdot g) : g \in \text{GL}(2,\mathbb{A}_\mathbf{Q})\}}$: it turns out that $V(f)$ is an irreducible unitary representation π_f of $\text{GL}(2,\mathbb{A}_\mathbf{Q})$ – cf. Gelbart for a justification of these statements. Such a representation splits as a completed tensor product of unitary representations of the p-adic Lie groups $\text{GL}(2,\mathbf{Q}_p)$:

$$\pi_f \cong \widehat{\bigotimes_{p \leq \infty}} \pi_{f,p}.$$

The local Langlands conjectures require that the $\pi_{f,p}$'s correspond to representations ρ_p of the Weil-Deligne group of $\mathbf{Q}_p$. The local/global compatibility we had in mind is, with $\mathcal{W}_{\mathbf{Q}_p} \subset G_{\mathbf{Q}_p}$ the absolute Weil group:

$$\rho_f\big|_{\mathcal{W}_{\mathbf{Q}_p}} = \rho_p\big|_{\mathcal{W}_{\mathbf{Q}_p}}.$$

5.3. **Consequence.** Let f, ρ_f be as in the theorem; $\rho_f(G_\mathbf{Q})$ is a finite subgroup of $\text{GL}(2,\mathbf{C})$: so is $\overline{\rho_f(G_\mathbf{Q})}$ in $\text{PGL}(2,\mathbf{C})$, which is thus conjugate to one of the following: $\mathbf{Z}/n\mathbf{Z}, D_{2n}, \mathfrak{A}_5, \mathfrak{A}_4, \mathfrak{S}_4$. The first is ruled out, because it's commutative (ρ_f would be a sum of two characters in this case); it is an interesting fact, not too difficult to establish, that the dihedral groups are

15

obtained precisely from modular forms induced from characters on quadratic extensions, as explained in section 2.9. One calls a form f *exotic* if $\overline{\rho_f(G_{\mathbf{Q}})}$ is isomorphic to one of $\mathfrak{A}_5, \mathfrak{A}_4, \mathfrak{S}_4$. For $\mathcal{B}_1^{new}(q, \chi)$ a basis of $\mathcal{S}_1^{new}(q, \chi)$, with forms chosen to be Hecke eigenforms, let

$$s^{Exotic}(q, \chi) = |\{f \in \mathcal{B}_1^{new}(q, \chi) : f \text{ is exotic }\}|.$$

This number is easily seen to be independent of the choice of $\mathcal{B}_1^{new}(q, \chi)$.

CONJECTURE 5.1. *With the notations above:*

$$s^{Exotic}(q, \chi) \ll_\varepsilon q^\varepsilon$$

As the conjecture, and its weakened forms, are statements uniform in q, it is important to see what happens when q varies. For example, if it is true that for q fixed, the number of Hecke eigenvalues of dihedral forms is bounded (they are sums of two roots of unity of order determined by q), when q varies, the number of possibilities grows with q. But, one can expect that for exotic forms, one can get a uniform bound on the number of Hecke eigenvalues.

Indeed, suppose that (q, χ) varies among pairs of integers and Dirichlet characters having order *bounded* by M. Let f be an exotic primitive (new)form. The image of $\rho_f(G_{\mathbf{Q}})$ in $\text{PGL}(2, \mathbf{C})$ has order bounded by 120; so the eigenvalues of $\rho(F_p)$ modulo the center are roots of unity of order less than 120, so the eigenvalues of $\rho(F_p)$ itself are roots of unity of order less than $120M$; so the $a_f(p)$, when p varies among the primes, is a sum of two such complex numbers. Let S denote this finite set. As a consequence, the polynomial $P(X) = \prod_{s \in S}(X - s) + A$ satisfies, for any prime p, any level q, and any exotic form f:

$$P(a_p(f)) = A$$

Using the Hecke-type relation at good primes:

$$a_p(f)^2 = a_{p^2}(f) + \chi(p)$$

expanding the product giving $P(a_p(f))$, and choosing A large enough, one gets a *linear* relation of the type:

$$\sum_{k=1}^{|S|} \mu_k(\chi) a_{p^k}(f) = B(\chi)$$

valid for any q, χ, and any good prime p, with bounded (when χ, q vary) complex numbers $\mu_k(\chi)$, $|B(\chi)| \geq 1$. This rough linear relation is sufficient to get a power saving as a first attempt to prove the (still unsolved) conjecture. The point is that one can suppress the assumption on the boundedness of the order of χ to get a linear relation among the $a_{p^k}(f)$, and even get an explicit (small) length, but this requires a careful and difficult analysis of the possible lifts from $\text{PGL}(2)$ to $\text{GL}(2)$. This linear relation depends on

the type of the form: in the case of icosahedral forms f (i.e. $\overline{\rho_f(G_\mathbf{Q})} \simeq \mathfrak{A}_5$) one has for example:

$$\bar{\chi}^6(p)a_{p^{12}}(f) - \bar{\chi}^4(p)a_{p^8}(f) - \bar{\chi}(p)a_{p^2}(f) = 1$$

The other relations are similar, but involve smaller powers of p, and so the bound obtained is better – cf. section 7.

6. Proof of the Deligne-Serre theorem.

The reader should refer to the original paper, and Serre's overview in the Durham volume. We will content ourselves to recall the main steps. The starting point is the following theorem (due to Deligne *et al.*):

THEOREM 6.1. *Let f be a weight $k \geq 2$ cuspidal modular form, with neben-typus χ, level q, Hecke eigenvector at the good primes with eigenvalues $\{a_p(f)\}_{(p,q)=1}$. Let K be a number field containing $\mathbf{Q}(f)$, and let λ be a finite prime of K of residue characteristic ℓ. Then there exists an irreducible λ-adic Galois representation, unique up to isomorphism, $\rho_f : G_\mathbf{Q} \longrightarrow \mathrm{GL}(2, K_\lambda)$, unramified away from $q\ell$, such that at any prime p not dividing $q\ell$:*

$$(5) \qquad\qquad \mathrm{Tr}(\rho_f(F_p)) = a_f(p)$$

$$(6) \qquad\qquad \det(\rho_f(F_p)) = p^{k-1}\chi(p)$$

Once Deligne-Serre is proven, the above is true for weight one as well, by choosing an algebraic embedding $\imath : K_\lambda \hookrightarrow \mathbf{C}$. The uniqueness statement follows from Tchebotarev as we saw in section 4 – this is the reason why one does not state the compatibility using the L-factors, which depend on the algebraic embedding, but this could be fixed by requiring that $\imath_{|\mathbf{Q}(f)} = \mathrm{id}_{\mathbf{Q}(f)}$.

6.1. **Step 1: Reduction modulo a prime.** Let's consider the following preliminary result:

THEOREM 6.2. *Let f be a weight $k \geq 1$ cuspidal modular form, with neben-typus χ, level q, Hecke eigenvector at the good primes with eigenvalues $\{a_p(f)\}_{(p,q)=1}$. Let K be a number field containing $\mathbf{Q}(f)$, and let λ be a finite prime of K of residue characteristic ℓ and let $k_f = \mathbf{F}_\ell(a_f(p), \chi(p); p \nmid q)$. Then there exists a semi-simple Galois representation, unique up to isomorphism, $\overline{\rho_f} : G_\mathbf{Q} \longrightarrow \mathrm{GL}(2, k_f)$, unramified away from $q\ell$, such that at any prime p not dividing $q\ell$:*

$$(7) \qquad\qquad \mathrm{Tr}(\overline{\rho_f}(F_p)) \equiv a_f(p)\,\mathrm{mod}(\lambda)$$

$$(8) \qquad\qquad \det(\overline{\rho_f}(F_p)) \equiv p^{k-1}\chi(p)\,\mathrm{mod}(\lambda).$$

One can strengthen the theorem, as it is actually only necessary that f be a Hecke eigenform modulo λ: see Deligne-Serre.

17

(1) If f is a weight one form, multiplication of f by a weight m $\mathrm{SL}(2,\mathbf{Z})$ Eisentein series produces a weight $m+1$ modular form. By considering the normalized Eisenstein series:

$$E_m(z) = 1 - \frac{b_m}{2m}\sum_{n=1}^{\infty}\sigma_{m-1}(n)\mathrm{e}(nz),$$

the Bernouilli numbers b_m satisfy congruence relations (Clausen-von Staudt theorem): $\ell b_m \equiv -1\,\mathrm{mod}(\ell)$ if $(\ell-1)|m$ (cf. Borevich-Shafarevitch, chap. 5 §8), so $fE_m \equiv f\,\mathrm{mod}(\lambda)$ for such a m.

(2) Re-establish good Hecke behaviour: there exists a weight $m+1$ modular form f', eigenform at good primes, defined over some extension K'/K, such that $a_p(f') \equiv a_p(f)\,\mathrm{mod}(\lambda')$ for some $\lambda'|\lambda$, and $p \nmid \ell q$.

(3) One can apply Deligne's theorem 6.1 to this f', getting a representation ρ' with values in $\mathrm{GL}(2, K'_{\lambda'})$.

(4) $\rho'(G_{\mathbf{Q}})$ is a compact subgroup of $\mathrm{GL}(2, K'_{\lambda'})$, which one can suppose to be contained in $\mathrm{GL}(2, \mathcal{O}'_{\lambda'})$, after possible conjugation. So one can reduce ρ' modulo λ', and get $\overline{\rho'}$ with values in $\mathrm{GL}(2, \mathcal{O}'_{\lambda'}/\lambda')$. Note that so far we preserved the congruences between the initial f and the Galois representations,

(5) but $\overline{\rho'}$ may not be semi-simple: taking its semi-simplification (cf. 4.2), one gets a semi-simple representation satisfying the congruences required in the theorem.

(6) One can reduce the field of definition from $\mathcal{O}'_{\lambda'}/\lambda'$ to k_f, because the characteristic polynomials of all the $\overline{\rho'}(g)$ have their coefficients in k_f by definition of k_f (note that Tchebotarev theorem has been used again): more precisely, $\overline{\rho'}$ is isomorphic to a representation $\overline{\rho_f}: G_{\mathbf{Q}} \longrightarrow \mathrm{GL}(2, k_f) \subset \mathrm{GL}(2, \mathcal{O}'_{\lambda'}/\lambda')$ (cf. Weintraub, Theorem 6.5, §3.6).

The unramifiedness is of course evident. One has lost irreducibility of the representations by reducing them. The study of (odd) Galois representations with values in finite fields led Serre to conjecture that they all arise as described above, which was proven recently by Khare *et al.*.

6.2. **Step 2: Bounding the images of the reductions.** In step one, one reduced Galois representations modulo a prime. To get a complex Galois representation, one needs to lift them back! This process is not trivial, but one can do it in the following situation:

PROPOSITION 6.1. *Let G be a finite group, ℓ be a prime not dividing $|G|$ and $\overline{\rho_\ell}: G \longrightarrow \mathrm{GL}(n, \mathbf{F}_\ell)$ a representation. Suppose there exists a number field K containing the roots of unity of order $|G|$, and a prime ideal λ for which $\mathcal{O}_K/\lambda = \mathbf{F}_\ell$. Then there exists a representation $\rho: G \longrightarrow \mathrm{GL}(n, \mathcal{O}_K)$ whose reduction modulo λ is isomorphic to $\overline{\rho_\ell}$.*

Cf. Weintraub, end of §3.6, for a proof.

So, if $K \supset \mathbf{Q}(f)$ is "sufficiently large", one decomposes $\overline{\rho_f}$ of theorem 6.2 into $G_\mathbf{Q} \longrightarrow \overline{\rho_f}(G_\mathbf{Q})$ followed by $\overline{\rho_f}(G_\mathbf{Q}) \hookrightarrow \mathrm{GL}(2, k_f)$, and would like to lift the second to $\mathrm{GL}(2, K)$. The result we quoted tells us this is possible, if we can assume that the characteristic of k_f is coprime to $|\overline{\rho_f}(G_\mathbf{Q})|$, at least for sufficiently many primes λ of K. This is indeed the case, and one proceeds as follows.

Let $K \supset \mathbf{Q}(f)$ be a finite Galois extension of $\mathbf{Q}$; let $\mathscr{S}$ be the set of primes of $\mathbf{Q}$, totally split in K. It is known that $\mathrm{Dens}(\mathscr{S}) = 1/[K : \mathbf{Q}]$, so $\mathscr{S}$ is infinite, and we have some flexibility in the choice of K. For $\ell \in \mathscr{S}$, theorem 6.2 gives a semi-simple representation $\overline{\rho_\ell} : G_\mathbf{Q} \longrightarrow \mathrm{GL}(2, \mathbf{F}_\ell)$. As a consequence, $G_\ell := \overline{\rho_\ell}(G_\mathbf{Q})$ is a semi-simple subgroup of $\mathrm{GL}(2, \mathbf{F}_\ell)$.

The key result is the following:

PROPOSITION 6.2. *With this notation, one has:*

$$\sup_{\ell \in \mathscr{S}} |G_\ell| < \infty$$

SKETCH OF PROOF: Recall that in section 3.6, we proved the existence of a set of primes $\mathscr{P}_\eta$, for any positive η, having density less than η, and a finite set of complex numbers Y_η such that:

$$\forall p \notin \mathscr{P}_\eta, a_p(f) \in Y_\eta.$$

Let $H_{\ell, \eta}$ be the set containing all the conjacy classes $\{\overline{\rho_\ell}(F_p)\}_{p \notin \mathscr{P}_\eta}$. By Tchebotarev, one has:

$$\mathrm{Dens}(\mathscr{P}_\eta^c) = \frac{|H_{\ell, \eta}|}{|G_\ell|} \geq 1 - \eta.$$

In other words, $H_{\ell, \eta}$ is a subset "large" inside G_ℓ, whose set of characteristic polynomials is small (less than $|Y_\eta|$, but in any case independent of ℓ). As the semi-simple subgroups of $\mathrm{GL}(2, \mathbf{F}_\ell)$ can be classified, one checks case by case the uniform bound claimed in the proposition: cf. Deligne-Serre, proposition 7.2, for the details.

6.3. **Step 3: Lift to** $\mathrm{GL}(2, \mathbf{C})$. One is now in position to apply proposition 6.1. Indeed, let $M = \sup_{\ell \in \mathscr{S}} |G_\ell|$: by adjoining them if necessary, on can suppose that K contains the M-th roots of unity; and let

$$\mathscr{S}' = \{p \text{ prime } : p \text{ totally split in K and } p > M\}.$$

Proposition 6.1 produces, for each $\ell \in \mathscr{S}'$, a representation $\rho_\ell : G_\mathbf{Q} \longrightarrow \mathrm{GL}(2, \mathcal{O}_K)$ whose reduction modulo a prime λ dividing ℓ is $\overline{\rho_\ell}$ (ℓ being totally split in K, any prime ideal λ dividing ℓ satisfies $\mathcal{O}_K/\lambda \cong \mathbf{F}_\ell$).

The last point to be careful with is the effect of this lifting on the congruences as stated in theorem 6.2. One can argue as follows. Consider the finite set of polynomials:

$$\mathcal{P}_\mathcal{M} = \{P(X) = (1 - \alpha X)(1 - \beta X) : \alpha, \beta \text{ } M\text{-th roots of unity}\}$$

19

(1) For each prime p, and $\ell \in \mathscr{S}', \ell \neq p$ and $\lambda | \ell$, the theorem 6.2 implies that:

$$\exists R_\ell \in \mathcal{P}_\mathcal{M} \; : \; 1 - a_p(f)X + \chi(p)X^2 \equiv R_\ell(X) \bmod(\lambda).$$

As $\mathcal{P}_\mathcal{M}$ is finite, the same polynomial works for infinitely many ℓ, so $1 - a_p(f)X + X^2$ itself is in $\mathcal{P}_\mathcal{M}$. Note that at this point, the Ramanujan conjecture is proven.

(2) As the set $\mathcal{P}_M$ is finite, one can suppose that $\forall P, Q \in \mathcal{P}_M ; P \neq Q \Rightarrow P \not\equiv Q \bmod(\lambda)$ (after removing a finite set of unsatisfying primes ℓ).

(3) Fix a prime ℓ not dividing the level q of the form f, and let $p \neq \ell$ as above. The lift ρ_ℓ to $\mathrm{GL}(2, \mathcal{O}_K)$ is unramified outside $q\ell$, and the characteristic polynomial of $\rho_\ell(F_p)$ is in $\mathcal{P}_\mathcal{M}$: this is because by construction $|\rho_\ell(G_\mathbf{Q})| = |G_\ell|$, so $\rho_\ell(F_p)$ has order less than M. On the other hand, the polynomial $1 - a_p(f)X + \chi(p)X^2$ is in $\mathcal{P}_\mathcal{M}$ as well, and as ρ_ℓ lifts $\overline{\rho_\ell}$, the two polynomials are congruent modulo λ: by the choice of ℓ, they are hence equal, so one has:

$$\forall p \nmid q\ell, \det(Id - X\rho_\ell(F_p)) = 1 + a_p(f)X + \chi(p)X^2.$$

(4) One has the same conclusion for another choice ℓ' (by replacing ℓ by ℓ').

(5) This means that the two representations $\rho_\ell, \rho_{\ell'} : G_\mathbf{Q} \longrightarrow \mathrm{GL}(2, \mathbf{C})$ are isomorphic as $\mathbf{C}$-representations, because of proposition 4.2 (after choosing an embedding $K \hookrightarrow \mathbf{C}$), that actually all the ρ_ℓ's are unramified outside q, and that the relation:

$$\forall p \nmid q, \det(Id - X\rho_\ell(F_p)) = 1 + a_p(f)X + \chi(p)X^2$$

holds. After this choice, one renames ρ_ℓ into ρ_f, and the construction is finished.

One should note that the finiteness of $\rho_f(G_\mathbf{Q})$ is a consequence of the construction; and that one gets the continuity of ρ_f from the finiteness of the image!

6.4. **Step 4: Irreducibility.** Reductio ad absurdum: one doesn't construct a stable subspace, but infers that reducibility would contradict the bound given on the Rankin-Selberg convolution seen in 3.4. Indeed, if ρ_f were not irreducible, by semi-simplicity (automatic over $\mathbf{C}$), one could find two one-dimensional stable subspaces of $\mathbf{C}^2$, on which $G_\mathbf{Q}$ would act by characters χ_1, χ_2, which by class field theory one can view as Dirichlet characters:

$$\rho_f \cong \chi_1 \oplus \chi_2$$

So $a_p(f) = \chi_1(p) + \chi_2(p)$, and $\chi(p) = \det(\rho_f(F_p)) = \chi_1(p)\chi_2(p)$, which proves by the way that $\chi_1 \neq \chi_2$ (else $\chi(-1) = 1$ and $f = 0$). One gets

$$\sum_{p \nmid q} \frac{|a_p(f)|^2}{p^s} = 2 \sum_{p \nmid q} p^{-s} + \sum_{p \nmid q} \frac{\chi_1(p)\overline{\chi_2(p)}}{p^s} + \sum_{p \nmid q} \frac{\chi_2(p)\overline{\chi_1(p)}}{p^s}$$

20

It is well known that for any non-trivial Dirichlet character ψ, $\sum_n \psi(n) n^{-s}$ is holomorphic at $s = 1$, so the two last terms are bounded. This would imply that

$$\sum_{p \nmid q} \frac{|a_p(f)|^2}{p^s} = 2 \sum_{p \nmid q} p^{-s} = -2 \log(s-1) + \mathcal{O}_{s \to 1}(1).$$

This contradicts the result given in section 3.4, and proves therefore the irreducibility of ρ_f.

7. THE DIMENSION OF THE SPACE OF WEIGHT ONE MODULAR FORMS.

Let's go back to the statements we gave on the dimension of weight one forms in the conjectures 5.1 (or 2.1, given in the case of squarefree levels, in which case the main term is explicitly related to quadratic inductions). We wrote

$$\dim(\mathcal{S}_1^{new}(q, \chi)) = s^{Dih}(q, \chi) + s^{Exotic}(q, \chi)$$

and expect that for any $\varepsilon > 0$

$$s^{Exotic}(q, \chi) \ll_\varepsilon q^\varepsilon$$

Duke *et al.* [Du] proved the following:

PROPOSITION 7.1. *There exists a positive real number δ such that:*

$$s^{Exotic}(q, \chi) \ll q^{1-\delta}.$$

Explicitly, the above inequality holds for any $0 \leq \delta < \frac{1}{12}$.

PROOF:

7.1. The idea is the following: let $\{x_n\}_{n \geq 1}$ be a sequence of complex numbers. By positivity one has the following inequality:

$$\sum_{f \in \mathcal{B}_1^{new}(q, \chi)} \left| \sum_{n=1}^N x_n c_n(f) \right|^2 \leq \sum_{f \in \mathcal{B}_1(q, \chi)} \left| \sum_{n=1}^N x_n c_n(f) \right|^2$$

for any orthonormal basis $\mathcal{B}_1^{new}(q, \chi)$ of the space of newforms, completed into an orthonormal basis $\mathcal{B}_1(q, \chi)$ of the total space space of weight one forms $\mathcal{S}_1(q, \chi)$. To use Galois representations, one chooses $\mathcal{B}_1^{new}(q, \chi)$ to be made of Hecke eigenforms, normalized ($a_1(f) = 1$) so that the Hecke eigenvalues and Fourier coefficients agree. This normalization destroys the unitarity of the forms, so if $\mathcal{B}_1^{new}(q, \chi)$ denotes now a Hecke eigenbasis of primitive forms, one has to study

$$\sum_{f \in \mathcal{B}_1^{new}(q, \chi)} \frac{1}{\|f\|^2} \left| \sum_{n=1}^N x_n a_n(f) \right|^2.$$

21

7.2. One makes three packets in $\mathcal{B}_1^{new}(q, \chi)$: octahedral, icosahedral, tetrahedral according to the type of the projective representation $\overline{\rho_f}$ ($\mathfrak{S}_4, \mathfrak{A}_5, \mathfrak{A}_4$ respectively). For each type, one chooses the test vector by using a linear relation, seen in section 5.3, of the type

$$\sum_{k=1}^{|S|} \mu_k(\chi) a_{p^k}(f) = B(\chi)$$

The longest linear relation is obtained for icosahedral forms, namely:

(9) $$\bar{\chi}^6(p) a_{p^{12}}(f) - \bar{\chi}^4(p) a_{p^8}(f) - \bar{\chi}(p) a_{p^2}(f) = 1$$

for which one chooses

$$x_n = \begin{cases} \bar{\chi}^6(p) & \text{if } n = p^{12}, (p, q) = 1 \\ -\bar{\chi}^4(p) & \text{if } n = p^8, (p, q) = 1 \\ -\bar{\chi}(p) & \text{if } n = p^2, (p, q) = 1 \\ 0 & \text{otherwise} \end{cases}$$

For this choice of $\{x_n\}$, one has for any icosahedral forms

$$\left| \sum_{n=1}^N x_n a_n(f) \right|^2 = \left| \{p \text{ prime} ; (p, q) = 1 \text{ and } p^{12} \leq N\} \right|^2 \gg \left(\frac{N^{\frac{1}{12}}}{\log(q)} \right)^2$$

by the prime number theorem; so one gets

$$\sum_{f \in \mathcal{B}_1^{ico}(q,\chi)} \frac{1}{\|f\|^2} \left(\frac{N^{\frac{1}{12}}}{\log(q)} \right)^2 \ll \sum_{f \in \mathcal{B}_1^{ico}(q,\chi)} \frac{1}{\|f\|^2} \left| \sum_{n=1}^N x_n a_n(f) \right|^2$$

7.3. On the other hand, the right hand side is bounded above by the large sieve inequality, which gives here

$$\sum_{f \in \mathcal{B}_1^{ico}(q,\chi)} \frac{1}{\|f\|^2} \left| \sum_{n=1}^N x_n a_n(f) \right|^2 \ll \left(1 + \frac{N}{q} \right) \left(\frac{N^{\frac{1}{12}}}{\log(q)} \right)$$

7.4. The two last steps give

$$\sum_{f \in \mathcal{B}_1^{ico}(q,\chi)} \frac{1}{\|f\|^2} \ll \left(1 + \frac{N}{q} \right) \left(\frac{N^{\frac{1}{12}}}{\log(q)} \right)^{-1}$$

which is uniform in q, χ, N; one still has to choose N, which depends on the behaviour of $\|f\|$ for a primitive form of level q.

22

7.5. Such an estimate is classical: cf. [I] section 13.6, and [Mi] section 2.3. Indeed:

$$\|f\|^2 := \int_{\Gamma_0(q)\backslash\mathfrak{H}} |\sqrt{y}f(z)|^2 \frac{dxdy}{y^2}$$

$$\doteq \operatorname{res}_{s=1} \operatorname{vol}(\Gamma_0(q)\backslash\mathfrak{H}) \int_{\Gamma_0(q)\backslash\mathfrak{H}} |\sqrt{y}f(z)|^2 E_s(z) \frac{dxdy}{y^2}$$

where the implied constant in "$\doteq$" is absolute, and with the usual Eisenstein series

$$E_s(z) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(q)} \frac{y^s}{|cz+d|^{2s}}$$

By unfolding the Eisenstein series, one proves that:

$$\int_{\Gamma_0(q)\backslash\mathfrak{H}} |\sqrt{y}f(z)|^2 E_s(z) \frac{dxdy}{y^2} = L(s, f \times \bar{f})$$

On the other hand, the Rankin-Selberg L-series (or its adelic counterpart) factorizes as[†]

$$L(s, f \times \bar{f}) \doteq L(s, 1)L(s, f, \operatorname{Ad})$$

which shows that

$$\operatorname{res}_{s=1} L(s, f \times \bar{f}) \doteq L(1, f, \operatorname{Ad})$$

and, at this point, one uses the well-known bound (cf. [Mi], section 1.3)

$$L(s, f, \operatorname{Ad}) \ll_\varepsilon q^\varepsilon$$

This gives:

$$\|f\|^2 \ll_\varepsilon \operatorname{vol}(\Gamma_0(q)\backslash\mathfrak{H})q^\varepsilon \ll_\varepsilon q^{1+\varepsilon}$$

7.6. By plugging the last estimate in 7.4, one has

$$|\mathcal{B}_1^{ico}(q,\chi)| \ll_\varepsilon q^{1+\varepsilon}\left(1+\frac{N}{q}\right)\left(\frac{N^{\frac{1}{12}}}{\log(q)}\right)^{-1}$$

7.7. By choosing $N = q$, one gets the result for icosahedral forms, i.e.:

$$|\mathcal{B}_1^{ico}(q,\chi)| \ll_\varepsilon q^{1-\frac{1}{12}+\varepsilon}$$

One does the same for the other types of forms – the only difference is that the linear relation among the Hecke eigenvalues have shorter length, and give a better estimate (better towards the conjecture).

7.8. **Remark.** If one uses the very weak linear relation obtained in section 5.3, one gets the estimate – cf. 5.3 for the notations:

$$s^{Exotic}(q,\chi) \ll_\varepsilon q^{1-|S|^{-1}+\varepsilon}$$

which is still a power saving, and uses much less analysis as the one required to get the explicit identity (9).

[†]For the definition of the adjoint L-function $L(s, f, \operatorname{Ad})$, see [Rog] section 13 e.g.

7.9. **Remark.** It is striking that one can ameliorate the bound of proposition 7.1, without changing the input: Michel-Venkatesh [MV] noted that the use of Kuznetsov trace formula, instead of the large sieve inequality, gives a bound in $q^{6/7}$ – instead of $q^{11/12}$ here. Here is how they proceed: embed the space of weight one modular forms into the space of weight one Maass cuspforms, as explained in section 3.2, and complete the basis $\mathcal{B}_1^{new}(q,\chi)$ into a basis $\mathcal{M}_1^{new}(q,\chi)$ of Maass cuspforms. One gets, by choosing a test function Φ non-negative on the spectrum of the hyperbolic Laplacian:

$$\sum_{f\in\mathcal{B}_1^{new}(q,\chi)} \Phi(0)\left|\sum_{n=1}^N x_n c_n(f)\right|^2 \leq \sum_{f\in\mathcal{M}_1(q,\chi)} \Phi(t_f)\left|\sum_{n=1}^N x_n c_n(f)\right|^2$$

One still has the lower bound:

$$\sum_{f\in\mathcal{B}_1^{new}(q,\chi)} \left|\sum_{n=1}^N x_n c_n(f)\right|^2 \gg_\varepsilon \left(\frac{N^{\frac{1}{12}}}{\log(q)}\right)^2 q^{-1-\varepsilon} s^{Exotic}(q,\chi)$$

To get an upper bound, one expands the square, and uses the Kuznetsov formula, to get

$$\sum_{f\in\mathcal{M}_1(q,\chi)} \Phi(t_f)\left|\sum_{n=1}^N x_n c_n(f)\right|^2 \doteq \sum_{n=1}^N |x_n|^2$$

$$+ \sum_{n,m=1}^N x_n\overline{x_m} \sum_{c\equiv 0\,(q)} \frac{K\ell_\chi(n,m,q)}{|c|}\widehat{\Phi}\left(\frac{4\pi\sqrt{nm}}{|c|}\right)$$

$$\ll_{\varepsilon,\Phi} \sum_{n=1}^N |x_n|^2 + (qN)^\varepsilon \frac{\sqrt{N}}{q}\left(\sum_{n=1}^N |x_n|\right)^2$$

where the last estimate follows from Weil's bound on Kloosterman sums. By choosing the same test vector $\{x_n\}$ as before, one gets the bound

$$s^{Exotic}(q,\chi) \ll_\varepsilon q^{\frac{6}{7}+\varepsilon}$$

8. On the Fourier coefficients of weight one modular forms.

Our last application of Deligne-Serre theorem is an estimate of the number of non-vanishing Fourier coefficients of weight one modular forms. Once one can use the associated Galois representation, the argument is incredibly simple and short, and is a nice interaction between worlds of Galois representations and of the analytic theory of modular forms. The problem is the following:

PROBLEM 8.1. *Let $f \in \mathcal{S}_1^{new}(q,\chi)$. What is the behaviour, when $N \to \infty$, of the quantity $\frac{1}{N}\left|\left\{n \leq N : a_n(f) \neq 0\right\}\right|$?*

The idea is the following: one first studies the question over prime integers, for which the number $a_p(f)$ is related to the trace of a matrix; then one bootstraps the estimate to get the density over the integers. The last section of Deligne-Serre article discusses the problem. Let $f \in \mathcal{S}_1^{new}(q, \chi)$, ρ_f its associated Galois representation given by theorem 5.1, and let $E_f/\mathbf{Q}$ be the finite Galois extension whose Galois group is $G_{\mathbf{Q}}/\ker(\rho_f)$.

8.1.   For any prime p, one has $a_p(f) = \mathrm{Tr}(\rho_f(F_p))$. Let:

$$Z(\rho_f) = \{g \in Gal(E_f/\mathbf{Q}) \ : \ \mathrm{Tr}(\rho_f(g)) = 0\}$$

One can apply Tchebotarev's theorem, to get:

$$\mathrm{Dens}\left(\left\{p \text{ prime } : \ a_p(f) = 0\right\}\right) = \mathrm{vol}(Z(\rho_f)) = \beta > 0$$

The number β is non-zero, because the (class of the) complex conjugation belongs to $Z(\rho_f)$.

8.2.   One uses Wirsing's theorem:

THEOREM 8.1. *Let ω be a non-negative multiplicative function, such that:*

(10) $$\sum_{p \leq N} \omega(p) \log(p) = \Big(\alpha + o_{N \to \infty}(1)\Big) N$$

(11) $$\exists \gamma_1, \gamma_2 > 0 : \forall p \text{ prime and } k \geq 0, f(p^k) \leq \gamma_1 \gamma_2^k$$

Then one has:

$$\sum_{n \leq N} \omega(n) = \Big(1 + o_{N \to \infty}(1)\Big) \cdot \frac{N}{\log(N)} \cdot \frac{\exp(-\gamma\alpha)}{\Gamma(\alpha)} \cdot \prod_{p \leq N} \left(\sum_{k=0}^{\infty} \frac{\omega(p^k)}{p^k}\right)$$

where γ denotes the Euler constant.

If one applies the theorem for $\omega(n) = 1$ if $a_f(n) \neq 0$, $\omega(n) = 0$ otherwise, one gets:

$$\left|\left\{n \leq N \ : \ a_n(f) \neq 0\right\}\right| = \Big(C + o_{N \to \infty}(1)\Big) \times \frac{N}{\log(N)^{\beta}}$$

for C some real number (depending on f), as the product over the primes has an asymtotic contribution for the primes such that $a_f(p) \neq 0$, which is a set of density $\alpha = 1 - \beta$.

8.3. **Remark.** One can rephrase these results as follows: for a fixed primitive eigenform of weight one f, $h : \mathbf{R} \to \mathbf{C}$ a bounded continuous function:

$$\frac{1}{N} \sum_{n=1}^{N} h(a_n(f)) \underset{N \to \infty}{\longrightarrow} h(0)$$

25

For a fixed form f, the set of possible values of the $a_p(f)$ is finite: let $\{a_i\}_i$ denote this set, and let β_i be the density of primes p for which $a_p(f) = a_i$. Then, one has for any continuous function $h : [-2, 2] \to \mathbf{C}$:

$$(12) \qquad \frac{\log(N)}{N} \sum_{p=2}^{N} h(a_p(f)) \underset{N \to \infty}{\longrightarrow} \sum_i \alpha_i h(a_i)$$

which shows that the sequence of the probability measures associated to the Hecke eigenvalues weakly converges to the discrete measure appearing on the right hand side. This situation is diametrically opposed to what is expected to hold for higher weight modular forms:

8.4. **The Sato-Tate conjecture.**

CONJECTURE 8.1 (Sato-Tate). *Let f be a (non-CM) primitive newform of weight $k \geq 2$, level q. Let $\{a_p(f)\}_{p \text{ prime}}$ denote its Hecke eigenvalues. For any continuous function $h : [-2, 2] \to \mathbf{C}$, one has:*

$$(13) \qquad \frac{\log(N)}{N} \sum_{p=2}^{N} h\left(\frac{a_p(f)}{p^{\frac{k-1}{2}}}\right) \underset{N \to \infty}{\longrightarrow} \frac{1}{\pi} \int_{-2}^{2} h(u)\sqrt{1 - \frac{u^2}{4}}\,du.$$

Cf. Ribet [R1] for a definition of non-CM in this setting. The difference between the limit distributions appearing in (12) (discrete atomic probability), and (13) (given by a density), is another instance of the dichotomy weight one/weight greater than one. Another difference is that Sato-Tate conjecture has been proven (recently) for some non-CM forms of weight two, corresponding to elliptic curves without complex multiplication. If one takes Sato-Tate conjecture for granted, one deduces easily that:

$$\mathrm{Dens}\left(\left\{p \text{ prime } : a_p(f) = 0\right\}\right) = 0$$

I would like to sketch here an unconditional proof of the above equality, as it illustrates the power of Galois representations – when one has them at one's disposal!

So let f be a (cuspidal) primitive form of weight $k \geq 2$ and level q. If K is a number field containing $\mathbf{Q}(f)$, and $\lambda | \ell$ a prime of K of residue characteristic ℓ not dividing the level, theorem 6.1 gives a λ-adic representation:

$$\rho_{f,\lambda} : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}(2, K_\lambda)$$

One can assume that $\rho_{f,\lambda}(G_{\mathbf{Q}}) \subset \mathrm{GL}(2, \mathcal{O}_\lambda)$ (after conjugating if necessary). Let $\overline{\rho_{f,\lambda}}$ be the (semi-simplification of the) reduction modulo λ of $\rho_{f,\lambda}$. Then, trivially, one has:

$$Z(f) := \left\{p \text{ prime} : a_p(f) = 0\right\} \subset Z(f, \lambda) := \left\{p \text{ prime} : a_p(f) \equiv 0 \,\mathrm{mod}(\lambda)\right\}$$

Let $E_\lambda/\mathbf{Q}$ be the finite Galois extension corresponding to the sugbroup $\ker(\overline{\rho_{f,\lambda}}) \subset G_\mathbf{Q}$; by Tchebotarev, one has:

$$\mathrm{Dens}\left(Z(f,\lambda)\right) = \frac{|\{g \in \overline{\rho_{f,\lambda}}(G_\mathbf{Q}) \, : \, \mathrm{Tr}(g) \equiv 0\}|}{|\overline{\rho_{f,\lambda}}(G_\mathbf{Q})|}$$

Suppose for the moment that one knows that $\rho_{f,\lambda}(G_\mathbf{Q}) = \mathrm{GL}(2,\mathcal{O}_\lambda)$ (which happens for almost all primes in the case of a non-CM elliptic curve). Then, by letting $q_\ell = |\mathcal{O}_\lambda/\lambda|$, one has:

$$|\overline{\rho_{f,\lambda}}(G_\mathbf{Q})| = (q_\ell^2-1)(q_\ell^2-q_\ell)\,;\, |\{g \in \overline{\rho_{f,\lambda}}(G_\mathbf{Q}) \, : \, \mathrm{Tr}(g) \equiv 0\}| = (q_\ell^2-1)(q_\ell-1)$$

which gives:

$$\lim_{\ell \to \infty} \mathrm{Dens}\left(Z(f,\lambda)\right) = 0$$

Unfortunately, the images of Galois representations corresponding to modular forms are not always open. Ribet studied in different works this problem, and gave in [R2] a general result, valid for non-CM modular forms. For our purposes, it is sufficient to say that for *totally split* primes ℓ,

$$\rho_{f,\lambda}(G_\mathbf{Q}) = \{g \in \mathrm{GL}(2,\mathbf{Z}_\ell) \, : \, \det(g) \in \mathbf{Z}_\ell^{\times(k-1)}\}$$

Here $\mathbf{Z}_\ell^{\times(k-1)}$ is the sugbroup of $k-1$ powers in $\mathbf{Z}_\ell^\times$. The subgroup $\overline{\rho_{f,\lambda}}(G_\mathbf{Q})$ is open and cofinite inside $\mathrm{GL}(2,\mathbf{Z}_\ell)$, with index independent of ℓ, and so one can argue as we did above! Indeed, $[\mathbf{Z}_\ell^\times : \mathbf{Z}_\ell^{\times(k-1)}] = \frac{|\mu_{k-1}(\mathbf{Q}_p)|}{|k-1|_p}$ stays bounded when $\ell \to \infty$ among totally split primes; so $[\mathrm{GL}(2,\mathbf{Z}_\ell) : \rho_{f,\lambda}(G_\mathbf{Q})]$ does as well (the two last indexes are equal to each other); as a consequence $[\mathrm{GL}(2,\mathbf{F}_\ell) : \overline{\rho_{f,\lambda}(G_\mathbf{Q})}]$ $(\leq [\mathrm{GL}(2,\mathbf{Z}_\ell) : \rho_{f,\lambda}(G_\mathbf{Q})])$ is uniformly bounded as the totally split primes ℓ go to infinity, and we have:

$$\frac{|\{g \in \overline{\rho_{f,\lambda}}(G_\mathbf{Q}) \, : \, \mathrm{Tr}(g) \equiv 0\}|}{|\overline{\rho_{f,\lambda}}(G_\mathbf{Q})|} \ll_{k,f} \frac{(\ell^2-1)(\ell-1)}{(\ell^2-1)(\ell^2-\ell)}$$

This shows unconditionally, that:

$$\mathrm{Dens}\left(\left\{p \text{ prime } ; \, a_p(f) = 0\right\}\right) = 0$$

in the case of a non-CM modular form of weight $k \geq 2$.

8.5. **Remark.** We have not dealt with forms having complex multiplication, but one expects a limit density; for CM elliptic curves, the limit distribution is the sum of a measure with density on $[-2,2]$ (pushforward of the uniform measure on the circle) and a Dirac at zero (coming from inert primes), and this is unconditionally proven (cf. Murty in [CKM], Lecture 1, end of §2).

8.6. **A conjecture.** Let's suppose that q is a prime number here, and that χ_q is Legendre character. Let f be a primitive Hecke eigenform in $\mathcal{S}_1(q, \chi_q)$. We proved in (12) the existence of numbers $\{\alpha_i(f)\}_{1 \leq i \leq N_f}$ with sum 1, and of real numbers $\{a_i(f)\}_{1 \leq i \leq N_f}$ in $[-2, 2]$ such that:

$$\frac{1}{\pi(N)} \sum_{p=2}^{N} \varphi(a_p(f)) \xrightarrow[N \to \infty]{} \mu_f\Big(\varphi\Big) := \sum_{i=1}^{N_f} \alpha_i(f)\varphi\big(a_i(f)\big)$$

for any continuous function $\varphi : [-2, 2] \to \mathbf{C}$. The limit distribution μ_f is discrete, but one can wonder if one re-establishes a density after averaging over a basis of newforms.

CONJECTURE 8.2. *Let $\mathcal{B}_1(q)$ be the basis of primitive newforms of $\mathcal{S}_1(q, \chi_q)$. Then the sequence of probability measures*

$$\mu_q := \frac{1}{\dim\Big(\mathcal{S}_1(q, \chi_q)\Big)} \sum_{f \in \mathcal{B}_1(q)} \mu_f$$

converges weakly as the prime $q \equiv 3 \bmod(4)$ goes to infinity. More precisely, for any continuous function $\varphi : [-2, 2] \to \mathbf{C}$, one has:

$$\frac{1}{\dim\Big(\mathcal{S}_1(q, \chi_q)\Big)} \sum_{f \in \mathcal{B}_1(q)} \mu_f\Big(\varphi\Big) \xrightarrow[q \to \infty]{} \frac{\varphi(0)}{2} + \frac{1}{4\pi} \int_{-2}^{2} \varphi(x) \frac{dx}{\sqrt{1 - \frac{x^2}{4}}}$$

If conjecture 5.1 is true, or if $s_1^{Exotic}(q, \chi_q) = o_{q \to \infty}\Big(s_1^{Dih}(q, \chi_q)\Big)$, then conjecture 8.2 is plausible, as the only contribution comes from dihedral forms. Assuming that the torsion in the class groups of imaginary quadratic fields is "small" (cf. [EV] for the terminology), we can then show that the mean value of such class group characters are equidistributed in the unit circle, as the discriminant tends to infinity. The Dirac comes from inert primes: if f is dihedral, then the primes inert in $\mathbf{Q}(\sqrt{-q})$ form a set of density $1/2$, and they all have $a_p(f) = 0$.

An even more challenging question is the following ("horizontal" Sato-Tate). One fixes a prime p, and considers for a continuous function φ as above the expression

$$\mu_{p,q}(\varphi) := \frac{1}{\dim\Big(\mathcal{S}_1(q, \chi_q)\Big)} \sum_{f \in \mathcal{B}_1(q)} \varphi(a_p(f))$$

Serre studied the weak limit of this sequence of measures, when the weight is greater than two, as the level q tends to infinity, using Eichler-Selberg's trace formula (see section 29 of [KL]). It is plausible that in the harmonically degenerated case of weight one forms, there is no limit anymore; but it would be interesting to determine the weak cluster points, and it is possible

that the smoothed sequence $\widetilde{\mu}_{p,Q} := \frac{\log(Q)}{Q} \sum_{q \leq Q} \mu_{p,q}$ has a limit when the prime Q tends to infinity.

Finally, the depth aspect. Fix an imaginary quadratic field and consider a family of ray class characters whose conductors are unbounded, or similarly a family of ring-class characters associated to a decreasing family of orders: in any of these examples, when the characters induce a weight one modular form of a certain level, it is possible to compare the size of the dihedral dimension to the exotic dimension, and to look at the weak limits of the corresponding families of measures.

References

[B]   Billingsley, Patrick: Convergence of probability measures. John Wiley & Sons, Inc., New York-London-Sydney 1968 xii+253 pp.

[BH]   Bushnell, Colin J.; Henniart, Guy: The local Langlands conjecture for GL(2). Grundlehren der Mathematischen Wissenschaften, 335. Springer-Verlag, Berlin, 2006.

[Bu]   Bump, Daniel: Automorphic Forms and Representations, Cambridge Studies in Advanced Mathematics, 55, (1997).

[C]   Casselman, William: On some results of Atkin and Lehner. Math. Ann. 201 (1973), 301–314.

[CF]   Algebraic number theory. Edited by J. W. S. Cassels and A. Frhlich Academic Press, London; Thompson Book Co., Inc., Washington, D.C. 1967 xviii+366 pp.

[CKM]   Cogdell, James W.; Kim, Henry H.; Murty, M. Ram: Lectures on automorphic L-functions. Fields Institute Monographs, 20. American Mathematical Society, Providence, RI, 2004. xii+283 pp.

[CR]   Curtis, Charles W.; Reiner, Irving: Representation theory of finite groups and associative algebras. Pure and Applied Mathematics, Vol. XI Interscience Publishers, a division of John Wiley & Sons, New York-London 1962 xiv+685 pp.

[CSS]   Modular forms and Fermat's last theorem. Edited by Gary Cornell, Joseph H. Silverman and Glenn Stevens. Springer-Verlag, New York, 1997. xx+582 pp.

[DI]   Diamond, Fred; Im, John: Modular forms and modular curves. Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994), 39–133, CMS Conf. Proc., 17, Amer. Math. Soc., Providence, RI, 1995.

[DS]   Deligne, Pierre; Serre, Jean-Pierre: Formes modulaires de poids 1. Ann. Sci. cole Norm. Sup. (4) 7 (1974), 507–530 (1975).

[Du]   Duke, William: The dimension of the space of cusp forms of weight one. Internat. Math. Res. Notices 1995, no. 2, 99–109.

[EV]   Ellenberg, Jordan; Venkatesh, Akshay: Reflection principles and bounds for class group torsion. Int. Math. Res. Not. IMRN 2007, no. 1, Art. ID # rnm002.

[G]   Gelbart, Stephen S.: Automorphic forms on adele groups, Annals of Math. studies, 83, Princeton University Press (1975).

[H]   Husemöller, Dale: Elliptic curves. Second edition. With appendices by Otto Forster, Ruth Lawrence and Stefan Theisen. Graduate Texts in Mathematics, 111. Springer-Verlag, New York, 2004. xxii+487 pp.

[I] Iwaniec, Henryk: Topics in classical automorphic forms. Graduate Studies in Mathematics, 17. American Mathematical Society, Providence, RI, 1997. xii+259 pp.

[IK] H. Iwaniec, E. Kowalski: Analytic Number Theory, American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004. xii+615 pp.

[ILS] Iwaniec, Henryk; Luo, Wenzhi; Sarnak, Peter: Low lying zeros of families of L-functions. Inst. Hautes tudes Sci. Publ. Math. No. 91 (2000), 55–131 (2001).

[KL] Knightly, Andrew; Li, Charles: Traces of Hecke operators. Mathematical Surveys and Monographs, 133. American Mathematical Society, Providence, RI, 2006. x+378 pp.

[M] Miyake, Toshitsune: Modular forms. Translated from the Japanese by Yoshitaka Maeda. Springer-Verlag, Berlin, 1989. x+335 pp.

[Mi] Michel, Philippe: Analytic number theory and families of automorphic L-functions, IAS/PCMI lecture series.

[MV] Michel, Philippe; Venkatesh, Akshay: On the dimension of the space of cusp forms associated to 2-dimensional complex Galois representations. Int. Math. Res. Not. 2002, no. 38, 2021–2027.

[R1] Ribet, Kenneth A.: Galois representations attached to eigenforms with Nebentypus. Modular functions of one variable, V (Proc. Second Internat. Conf., Univ. Bonn, Bonn, 1976), pp. 17–51. Lecture Notes in Math., Vol. 601, Springer, Berlin, 1977.

[R2] Ribet, Kenneth A.: On l-adic representations attached to modular forms. II. Glasgow Math. J. 27 (1985), 185–194.

[Rog] Rogawski, Jonathan D.: Functoriality and the Artin conjecture. Representation theory and automorphic forms (Edinburgh, 1996), 331–353, Proc. Sympos. Pure Math., 61, Amer. Math. Soc., Providence, RI, 1997.

[Roh] Rohrlich, David E.: Elliptic curves and the Weil-Deligne group. Elliptic curves and related topics, 125–157, CRM Proc. Lecture Notes, 4, Amer. Math. Soc., Providence, RI, 1994.

[S1] Serre, Jean-Pierre: Modular forms of weight one and Galois representations. Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975), pp. 193–268. Academic Press, London, 1977.

[S2] Serre, Jean-Pierre: Abelian l-adic representations and elliptic curves. Research Notes in Mathematics, 7. A K Peters, Ltd., Wellesley, MA, 1998. 199 pp.

[T] Taylor, Richard: Galois representations. Ann. Fac. Sci. Toulouse Math. (6) 13 (2004), no. 1, 73–119.

[Ta] Tate, John: Number theoretic background. Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2, pp. 3–26, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979.

[W] Weintraub, Steven H.: Representation theory of finite groups: algebra and arithmetic. Graduate Studies in Mathematics, 59. American Mathematical Society, Providence, RI, 2003. x+212 pp.

Stanford University, Department of Mathematics, building 380, Stanford, California 94305, USA.

E-mail: `trotabas@math.stanford.edu`